



OIPE

#2

RAW SEQUENCE LISTING

DATE: 02/01/2002

PATENT APPLICATION: US/10/006,163

TIME: 11:47:54

Input Set : N:\Crif3\RULE60\10006163.raw

Output Set: N:\CRF3\02012002\J006163.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Lal, Preeti
- 6 Corley, Neil C.
- 8 (ii) TITLE OF INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE
- 10 (iii) NUMBER OF SEQUENCES: 3
- 12 (iv) CORRESPONDENCE ADDRESS:
- 13 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
- 14 (B) STREET: 3174 Porter Dr.
- 15 (C) CITY: Palo Alto
- 16 (D) STATE: CA
- 17 (E) COUNTRY: USA
- 18 (F) ZIP: 94304

20 (v) COMPUTER READABLE FORM:

- 21 (A) MEDIUM TYPE: Diskette
- 22 (B) COMPUTER: IBM Compatible
- 23 (C) OPERATING SYSTEM: DOS
- 24 (D) SOFTWARE: FastSEQ for Windows Version 2.0

ENTERED

26 (vi) CURRENT APPLICATION DATA:

- C--> 27 (A) APPLICATION NUMBER: US/10/006,163
- C--> 28 (B) FILING DATE: 04-Dec-2001
- 29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

- 32 (A) APPLICATION NUMBER: US/09/249,241
- 33 (B) FILING DATE:

36 (viii) ATTORNEY/AGENT INFORMATION:

- 37 (A) NAME: Billings, Lucy J.
- 38 (B) REGISTRATION NUMBER: 36,749
- 39 (C) REFERENCE/DOCKET NUMBER: PF-0475 US

41 (ix) TELECOMMUNICATION INFORMATION:

- 42 (A) TELEPHONE: 650-855-0555
- 43 (B) TELEFAX: 650-845-4166
- 44 (C) TELEX:

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 313 amino acids
- 51 (B) TYPE: amino acid
- 52 (C) STRANDEDNESS: single
- 53 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

- 56 (A) LIBRARY: PROSNOT01
- 57 (B) CLONE: 356351

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/006,163

DATE: 02/01/2002

TIME: 11:47:54

Input Set : N:\Crf3\RULE60\10006163.raw

Output Set: N:\CRF3\02012002\J006163.raw

```

61 Met Ala Ala Pro Met Asn Gly Gln Val Cys Val Val Thr Gly Ala Ser
62 1 5 10 15
63 Arg Gly Ile Gly Arg Gly Ile Ala Leu Gln Leu Cys Lys Ala Gly Ala
64 20 25 30
65 Thr Val Tyr Ile Thr Gly Arg His Leu Asp Thr Leu Arg Val Val Ala
66 35 40 45
67 Gln Glu Ala Gln Ser Leu Gly Gly Gln Cys Val Pro Val Val Cys Asp
68 50 55 60
69 Ser Ser Gln Glu Ser Glu Val Arg Thr Leu Phe Glu Gln Val Asp Arg
70 65 70 75 80
71 Glu Gln Gln Gly Arg Leu Asp Val Leu Val Asn Asn Ala Tyr Ala Gly
72 85 90 95
73 Val Gln Thr Ile Leu Asn Thr Arg Asn Lys Ala Phe Trp Glu Thr Pro
74 100 105 110
75 Ala Ser Met Trp Asp Asp Ile Asn Asn Val Gly Leu Arg Gly His Tyr
76 115 120 125
77 Phe Cys Ser Val Tyr Gly Ala Arg Leu Met Val Pro Ala Gly Gln Gly
78 130 135 140
79 Leu Ile Val Val Ile Ser Ser Pro Gly Ser Leu Gln Tyr Met Phe Asn
80 145 150 155 160
81 Val Pro Tyr Gly Val Gly Lys Ala Ala Cys Asp Lys Leu Ala Ala Asp
82 165 170 175
83 Cys Ala His Glu Leu Arg Arg His Gly Val Ser Cys Val Ser Leu Trp
84 180 185 190
85 Pro Gly Ile Val Gln Thr Glu Leu Lys Glu His Met Ala Lys Glu
86 195 200 205
87 Glu Val Leu Gln Asp Pro Val Leu Lys Gln Phe Lys Ser Ala Phe Ser
88 210 215 220
89 Ser Ala Glu Thr Thr Glu Leu Ser Gly Lys Cys Val Val Ala Leu Ala
90 225 230 235 240
91 Thr Asp Pro Asn Ile Leu Ser Leu Ser Gly Lys Val Leu Pro Ser Cys
92 245 250 255
93 Asp Leu Ala Arg Arg Tyr Gly Leu Arg Asp Val Asp Gly Arg Pro Val
94 260 265 270
95 Gln Asp Tyr Leu Ser Leu Ser Ser Val Leu Ser His Val Ser Gly Leu
96 275 280 285
97 Gly Trp Leu Ala Ser Tyr Leu Pro Ser Phe Leu Arg Val Pro Lys Trp
98 290 295 300
99 Ile Ile Ala Leu Tyr Thr Ser Lys Phe
100 305 310

```

102 (2) INFORMATION FOR SEQ ID NO: 2:

104 (i) SEQUENCE CHARACTERISTICS:

105 (A) LENGTH: 1387 base pairs

106 (B) TYPE: nucleic acid

107 (C) STRANDEDNESS: single

108 (D) TOPOLOGY: linear

110 (vii) IMMEDIATE SOURCE:

111 (A) LIBRARY: PROSN01

112 (B) CLONE: 356351

RAW SEQUENCE LISTING

DATE: 02/01/2002

PATENT APPLICATION: US/10/006,163

TIME: 11:47:54

Input Set : N:\Crif3\RULE60\10006163.raw

Output Set: N:\CRF3\02012002\J006163.raw

114 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

116	CTAACTTTGG	CCTGGGACTC	TGCCCCCTCTA	CCTCAGCACA	GAATCGCCCC	GGGTCCTACT	60
117	ACAGAATCAA	TCCTTGAACA	CTGCCTCCAC	GTCGCCGGCT	CAATCTGGGC	GAGAACCCAG	120
118	ACTTCCACCG	CAGCCCCGCA	ATCTGCAGAC	CTCAGCGGCA	GCGCAGGTGG	CAGACCTGCC	180
119	TCCTTTGCCT	GTGAGTCATG	GCAGCTCCCA	TGAATGGCCA	AGTGTGTGTG	GTGACTGGTG	240
120	CCTCCAGGGG	TATTGGCCGT	GGCATTGCCT	TGCAGCTCTG	CAAAGCAGGC	GCCACAGTTT	300
121	ACATCACTGG	CCGCCATCTG	GACACCCCTC	GCGTTGTTGC	TCAGGAGGCA	CAATCCCTCG	360
122	GGGGCCAATG	TGTGCCTGTG	GTGTGCGATT	CAAGCCAGGA	GAGTGAAGTG	CGAACGCTGT	420
123	TTGAGCAAGT	GGATCGGGAA	CAGCAAGGGC	GTCTAGATGT	GCTGGTCAAC	AATGCTTATG	480
124	CAGGGGTCCA	GACGATCCTG	AACACCAGGA	ATAAGGCATT	CTGGGAAACC	CCTGCCTCCA	540
125	TGTGGGATGA	TATCAACAAC	GTCGGACTCA	GAGGCCACTA	CTTTTGCTCA	GTGTATGGGG	600
126	CACGGCTGAT	GGTACCAGCT	GGCCAGGGGC	TCATCGTGGT	CATCTCCTCC	CCAGGAAGCC	660
127	TGCAGTATAT	GTTCAATGTC	CCCTATGGTG	TGGGCAAAGC	TGCGTGTGAC	AAGCTGGCTG	720
128	CTGACTGTGC	CCACGAGCTG	CGGCGCCATG	GGGTCAGCTG	TGTGTCTCTG	TGGCCGGGGA	780
129	TTGTGCAGAC	AGAAGTGTG	AAGGAGCATA	TGGCAAAGGA	GGAGGTCTCTG	CAGGATCCTG	840
130	TGTTGAAGCA	GTTCAAATCA	GCCTTCTCAT	CTGCAGAAAC	CACAGAATTG	AGTGGCAAAT	900
131	GTGTGGTGGC	TTTGGCAACA	GATCCCAATA	TCCTGAGCCT	GAGTGGTAAG	GTGCTGCCAT	960
132	CCTGTGACCT	TGCTCGACGC	TATGGCCTTC	GGGATGTGGA	CGGCCGCCCC	GTCCAAGACT	1020
133	ATTTGTCTTT	GAGCTCTGTT	CTCTCACACG	TGTCCGGCCT	GGGCTGGCTG	GCCTCCTACC	1080
134	TGCCCTCCTT	CCTCCGTGTG	CCCAAGTGGA	TTATTGCCCT	CTACACTAGC	AAGTTCTAAC	1140
135	CCTCCTGGTC	TGACACTACG	TCTCTGCTTG	TCTTCTCATT	TGGACTTGGT	GGTTCGTCTT	1200
136	GTCTCAGTGA	AACAGCAGCC	TTTCTTGTTT	ACCCATACCC	TTGATATGAA	GAGAAGCCCT	1260
137	CTGCTGTGTG	TCCGTGGTGA	GTTCTGGGGT	GCGCCTAGGT	CCCTTCTTTG	TGCCTTGGTT	1320
138	TTCTTTGTCC	TTCTTTTTAC	TTTTTGCCCT	AGTATTGAAA	AATGCTCTTG	GAGCTAATAA	1380
139	AAGTCTA						1387

141 (2) INFORMATION FOR SEQ ID NO: 3:

143 (i) SEQUENCE CHARACTERISTICS:

144 (A) LENGTH: 323 amino acids

145 (B) TYPE: amino acid

146 (C) STRANDEDNESS: single

147 (D) TOPOLOGY: linear

149 (vii) IMMEDIATE SOURCE:

150 (A) LIBRARY: GenBank

151 (B) CLONE: 2315796

153 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

155	Met	Gly	Val	Ile	Leu	Gln	Asp	Gln	Val	Ala	Leu	Val	Thr	Gly	Ala	Ser
156	1			5					10						15	
157	Arg	Gly	Ile	Gly	Arg	Gly	Ile	Ala	Leu	Gln	Leu	Gly	Glu	Ala	Gly	Ala
158				20					25					30		
159	Thr	Val	Tyr	Ile	Thr	Gly	Arg	Arg	Pro	Glu	Leu	Ser	Asp	Asn	Phe	Arg
160				35				40					45			
161	Leu	Gly	Leu	Pro	Ser	Leu	Asp	Tyr	Val	Ala	Lys	Glu	Ile	Thr	Ser	Arg
162				50			55					60				
163	Gly	Gly	Lys	Gly	Ile	Ala	Leu	Tyr	Val	Asp	His	Ser	Asn	Met	Thr	Glu
164	65					70				75					80	
165	Val	Lys	Phe	Leu	Phe	Glu	Lys	Ile	Lys	Glu	Asp	Glu	Glu	Gly	Lys	Leu
166					85					90				95		
167	Asp	Ile	Leu	Val	Asn	Asn	Val	Tyr	Asn	Ser	Leu	Gly	Lys	Ala	Thr	Glu
168				100					105					110		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/006,163

DATE: 02/01/2002

TIME: 11:47:54

Input Set :: N:\Crf3\RULE60\10006163.raw

Output Set: N:\CRF3\02012002\J006163.raw

```

169 Met Ile Gly Lys Thr Phe Phe Asp Gln Asp Pro Ser Phe Trp Asp Asp
170      115      120      125
171 Ile Asn Gly Val Gly Leu Arg Asn His Tyr Tyr Cys Ser Val Tyr Ala
172      130      135      140
173 Ala Arg Met Met Val Glu Arg Arg Lys Gly Leu Ile Val Asn Val Gly
174 145      150      155      160
175 Ser Leu Gly Gly Leu Lys Tyr Val Phe Asn Val Ala Tyr Gly Ala Gly
176      165      170      175
177 Lys Glu Ala Leu Ala Arg Met Ser Thr Asp Met Ala Val Glu Leu Asn
178      180      185      190
179 Pro Tyr Asn Val Cys Val Val Thr Leu Ile Pro Gly Pro Val Lys Thr
180      195      200      205
181 Glu Thr Ala Asn Arg Thr Ile Ile Asp Asp Ala Tyr Lys Met Ile Lys
182      210      215      220
183 Glu Asn Pro Glu Leu Glu Glu Phe Ile Lys Gly Glu Ser Thr Glu Tyr
184 225      230      235      240
185 Thr Gly Lys Ala Leu Ala Arg Leu Ala Met Asp Pro Gly Lys Leu Lys
186      245      250      255
187 Lys Ser Gly Lys Thr Leu Phe Thr Glu Asp Leu Ala Gln Lys Tyr Asp
188      260      265      270
189 Phe Ser Asp Lys His Gly Ala Gly Met Glu Pro Gln Asn Ile Arg Ser
190      275      280      285
191 Ile Arg Thr Ile Leu Gly Thr Met Gly Lys Glu Glu Val Ala Lys Tyr
192      290      295      300
193 Ile Pro Pro Gln Ile Lys Leu Pro Lys Trp Val Ile Trp Gln Ser Val
194 305      310      315      320
195 Asn Arg Phe

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/006,163

DATE: 02/01/2002

TIME: 11:47:55

Input Set : N:\Crf3\RULE60\10006163.raw

Output Set: N:\CRF3\02012002\J006163.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]